ABSTRACT

A system (230) performs speaker adaptation when performing speech recognition. The system (230) receives an audio segment and identifies the audio segment as a first audio segment or a subsequent audio segment associated with a speaker turn. The system (230) then decodes the audio segment to generate a transcription associated with the first audio segment when the audio segment is the first audio segment and estimates a transformation matrix based on the transcription associated with the first audio segment. The system (230) decodes the audio segment using the transformation matrix to generate a transcription associated with the subsequent audio segment when the audio segment is the subsequent audio segment.